Worcester Polytechnic Institute
DigitalCommons@WPI

Computer Science Faculty Publications

Department of Computer Science

1-1-1998

A transformation-based combination framework for approximate reasoning

Sergio A. Alvarez
Worcester Polytechnic Institute, alvarez@cs.wpi.edu

A transformation-based combination framework for approximate reasoning WPI-CS-TR-98-22

Sergio A. Alvarez

Department of Computer Science
Worcester Polytechnic Institute [1]
Worcester, MA 01609

Abstract

There are many contexts in which several quantitative measures that provide information about a given phenomenon are available and it is desired to combine these measures into a single measure that uses the information encoded in each of them. Examples include knowledge aggregation in knowledge-based systems [4], [16], lateralization measurement in neurobiology [2], [5], and relevance ranking in information retrieval. Mostly ad-hoc approaches are currently in use for this purpose in different domains. The objective of this paper is to introduce a rational framework that systematically provides families of combination operators for the integration of disparate measures in a variety of situations. Our approach uses a single canonical form to produce a multitude of different combination functions by choosing different geometric frames of reference in the space of measurement values. We show that previously used combination functions may be obtained through our approach in a natural way, that they may be easily modified and generalized for increased flexibility, and that new combination operators may be systematically generated. We provide a characterization of the differentiable combination functions that are expressible via conjugacy in terms of the canonical form and give an algorithm to construct an appropriate reference frame if one exists. We also address the asymptotic behavior of the combination functions produced by our framework when the number of source measures grows without bound.

[1]Portions of this paper are based on research carried out by the author at the Center for Nonlinear Analysis at Carnegie Mellon University. The author acknowledges partial support provided by the U.S. Army Research Office and the National Science Foundation. The author wishes to thank Bruce Buchanan for helpful discussions regarding knowledge revision in the expert system MYCIN.

Introduction

The issue of combination or aggregation of knowledge sources is central to many areas of applied science and engineering. Consider for example the problem of knowledge revision in belief systems. Various approaches to this problem in the presence of uncertainty are elegantly subsumed by the Shenoy-Shafer valuation network theory [16] in which a network of *valuations* encodes approximate knowledge about the joint values of collections of system variables and knowledge revision is reduced to the two basic operations of *marginalization* and *combination* of valuations. Specific forms of the combination functions are provided within formalisms such as Bayesian probability and Dempster-Shafer belief theory. In Bayesian probability the valuations are true probabilities and combination proceeds according to Bayes' rule. The simplest version of this combines probabilities p and q by measuring the probability of the union of the corresponding events assuming independence between these events:

$$f(p,q) = p + q - pq \tag{1}$$

In the Dempster-Shafer theory the valuations are so-called basic probability assignments and combination follows Dempster's rule [15]. More ad-hoc approaches have also been proposed, as in the framework of certainty factors introduced into rule-based expert systems by the creators of the medical diagnosis system MYCIN [4]. In this method the valuations are numbers between -1 and 1 called *certainty factors* which represent confidence levels about both facts and inference rules. The MYCIN combination function takes two certainty factors c_1 and c_2 of different signs and combines them into a single certainty factor c as follows:

$$c = \frac{c_1 + c_2}{1 + \min(c_1, c_2)} \tag{2}$$

Because of the constraint on the signs of the c_i, this measure may be described as a difference measure rather than a sum measure like that in Eq. 1. Such difference measures are required also in the area of anatomical and functional lateralization measurement in biology [12], [5]. For example, in studying a bihemispheric brain one is interested in assessing the degree of asymmetry of the patterns of organization and functionality of the system. If one has access to two individual measures representing the competence of each of the hemispheres on some task of interest, then one may seek to combine these measures into a single measure of the lateral dominance of one hemisphere over the other as regards the given task. Related examples include the measurement of directional asymmetry in experimental psychology [6], [10] and high energy physics [7]. Prior work in the above mentioned areas has tended to use simple, ad-hoc measures of directional asymmetry, such as the standard arithmetic difference of the given unilateral measures.

In the present paper we present a systematic approach for generating numerical combination functions and related difference measures. We will show that several previously used measures are subsumed by our framework and we will propose mechanisms that yield rational generalizations and modifications of these measures as well as completely new ones. Our framework is based on thinking of different measures as corresponding to the same canonical form viewed in different geometric frames on the space of measurement values. We consider measures of sum and difference type interchangeably by allowing sign changes in the arguments. For concreteness, we now phrase our fundamental postulate in terms of combination (sum) functions only.

Postulate (existence of a canonical form). ***Combination functions should reduce to the standard arithmetic sum in a suitably constructed frame.***

In words, given an admissible combination function $f : V \times V \to V$, there should exist a suitable choice of *frame transformation* β such that we have a commutative diagram as shown below, where $+$ denotes the usual arithmetic sum operator on the real line $\mathrm{R} = \beta(V)$:

$$\begin{array}{ccc} V \times V & \xrightarrow{\beta\times\beta} & \beta(V)\times\beta(V) \\ \downarrow{\scriptstyle f} & & \downarrow{\scriptstyle +} \\ V & \xrightarrow{\beta} & \beta(V) \end{array}$$

Equivalently, the combination function f should satisfy

$$f\left(\beta^{-1}(a), \beta^{-1}(b)\right) = \beta^{-1}(a+b) \tag{3}$$

At first sight, the above may seem like an odd requirement. In the present paper we aim to show that the canonical form postulate is not only natural, being satisfied by commonly used combination operators already in existence and slight variations of them, but also constitutes a powerful source of new combination operators. In particular, through this unification, our new framework based on the canonical form and frame transformations provides a much needed theoretical foundation for combination operators.

Example 1. Consider the case of the following simple probabilistic combination function which is often used for aggregation of measures of uncertainty in knowledge-based systems:

$$f(p, q) = p + q - pq \tag{4}$$

Analysis shows that one may rewrite the function of Eq. 1 in the form given in Eq. 3, where the transformation $\beta : [0, 1] \to [0, \infty]$ that defines the "normalizing frame" is given by

$$\beta(x) = \log\left(\frac{1}{1-x}\right) \tag{5}$$

Indeed, the inverse of the frame transformation is

$$\beta^{-1}(y) = 1 - e^{-y} \tag{6}$$

and by direct computation using Eq. 1 and Eq. 6 we confirm that Eq. 3 holds:

$$\begin{aligned} f(\beta^{-1}(p), \beta^{-1}(q)) &= \beta^{-1}(p) + \beta^{-1}(q) - \beta^{-1}(p)\beta^{-1}(q) \\ &= 1 - (1 - \beta^{-1}(p))(1 - \beta^{-1}(q)) \\ &= 1 - e^{-p}e^{-q} \\ &= \beta^{-1}(p+q) \end{aligned} \tag{7}$$

The present paper provides, as part of a coherent theory, a method that allows one to construct an appropriate frame transformation β as in Eq. 5 directly from the combination function f. Related results in the case of difference measures only were obtained in [1].

Overview of the paper

We begin the paper by presenting a set of axioms which state the properties required for a combination function to be admissible. No assumptions are made about the particular method used to construct the functions at this point. We then study the class of combination functions defined from the canonical form via frame transformations on the range of the valuations as in the above commutative diagram and Eq. 3. We determine what properties a frame transformation must satisfy in order for the associated combination function to be admissible. We show that admissible frame transformations may be described "microscopically" in terms of a Riemannian metric associated with the *subjective difference measure* obtained from the frame transformation. We give examples of admissible frame transformations and the combination functions obtained from them. We then present a general method to extract a suitable normalizing frame transformation directly from a given combination function as was done for the above example in Eqs. 5 and 7. We conclude by describing the asymptotic behavior of aggregate values obtained via our transformation-based framework in the presence of an unbounded number of sources of information.

1 Admissible combination and difference functions

In this brief section we give axioms for the binary operations that we are interested in studying. The basic notion is that of a *combination function*, which is a generally nonlinear function that aggregates two different measurements into a single one. The simplest possible combination function is the arithmetic operation of addition. Just as addition yields subtraction by changing the sign of one of the arguments, any combination function gives rise to a *difference measure* in the same way. We give equivalent axioms for both combination functions and difference measures. One version or the other will typically be more immediately useful in a given context. For example, in lateralization measurement in computational neurobiology [2] one uses difference measures, while in knowledge aggregation in knowledge-based systems it is more natural to use combination functions.

Definition 1.1. A function $\oplus : [0, +1] \times [0, +1] \to [0, +1]$ is an *admissible combination function* if and only if it satisfies the following axioms:

Commutativity

$p \oplus q = q \oplus p$

Monotonicity

$(\cdot) \oplus q$ is an increasing function for each q

Boundary values

$0 \oplus q = q, \;\; 1 \oplus q = 1$

Definition 1.2. We define the *subjective difference measure* $\ominus$ associated with the combination operator $\oplus$ to be the operator $\ominus$ defined as follows:

$$p \ominus q = p \oplus (-q) \tag{8}$$

The operator $\ominus$ is said to be *symmetric* if it satisfies

$$q \ominus p = -(p \ominus q) \tag{9}$$

It is clear that $\ominus$ is symmetric if and only if the associated combination operator $\oplus$ satisfies

Belief / disbelief symmetry

$(-p) \oplus (-q) = -(p \oplus q)$

Example 2. The probabilistic combination operator given in the Example that appears in the Introduction is admissible in the sense of the above definition. The commutativity property clearly holds for this operator. Also, by rewriting the operator in the form

$$p \oplus q = p(1-q) + q,$$

it becomes apparent that $p \oplus q$ increases as p increases if q is held fixed. Finally, the boundary values for the probabilistic combination operator are given by:

$$0 \oplus q = 0(1-q) + q = q, \quad 1 \oplus q = 1(1-q) + q = 1$$

This proves admissibility as claimed.

Example 3. The MYCIN combination function is admissible. Recall that this combination function is defined by:

$$p \oplus q = \frac{p+q}{1 + p \wedge q},$$

where $p \wedge q$ denotes the minimum of the two numbers p and q. The properties of commutativity and boundary values are easy to see. Verification of the monotonicity property is conceptually simple but requires an analysis by cases. Assume that q is fixed and that $1 \geq p' \geq p \geq 0$. We must show that $p' \oplus q \geq p \oplus q$. The difference $p' \oplus q \; - \; p \oplus q$ equals

$$p' \oplus q \; - \; p \oplus q = \frac{p'+q}{1+p' \wedge q} - \frac{p+q}{1+p \wedge q} = \frac{p' - p + (p \wedge q)(p'+q) - (p' \wedge q)(p+q)}{(1+p' \wedge q)(1+p \wedge q)} \tag{10}$$

- Case 1: $q \leq p \leq p'$. Then $p \wedge q = q = p' \wedge q$, and the right-hand side of Eq. 10 becomes
$$\begin{aligned}\frac{p' - p + (p \wedge q)(p' + q) - (p' \wedge q)(p + q)}{(1 + p' \wedge q)(1 + p \wedge q)} &= \frac{p' - p + q(p' + q) - q(p + q)}{(1 + q)^2} \\ &= \frac{(p' - p)(1 + q)}{(1 + q)^2} \geq 0\end{aligned}$$

- Case 2: $p \leq q \leq p'$. Then $p \wedge q = p$ and $p' \wedge q = q$, so in Eq. 10 we have
$$\begin{aligned}\frac{p' - p + (p \wedge q)(p' + q) - (p' \wedge q)(p + q)}{(1 + p' \wedge q)(1 + p \wedge q)} &= \frac{p' - p + p(p' + q) - q(p + q)}{(1 + p' \wedge q)(1 + p \wedge q)} \\ &= \frac{p' - p + pp' - q^2}{(1 + q)(1 + p)} \\ &\geq \frac{p' - p + pp' - (p')^2}{(1 + q)(1 + p)} \\ &= \frac{(p' - p)(1 - p')}{(1 + q)(1 + p)} \geq 0\end{aligned}$$

- Case 3: $p \leq p' \leq q$. Then $p \wedge q = p$ and $p' \wedge q = p'$, and Eq. 10 yields:
$$\begin{aligned}\frac{p' - p + p(p' + q) - p'(p + q)}{(1 + p' \wedge q)(1 + p \wedge q)} &= \frac{p' - p + pq - p'q}{(1 + p')(1 + p)} \\ &= \frac{(p' - p)(1 - q)}{(1 + p')(1 + p)} \geq 0\end{aligned}$$

This concludes the verification of the monotonicity property and thus establishes that the MYCIN combination function is admissible in the sense of Definition 1.1.

In the next section we develop a framework that incorporates combination functions similar to those considered in the preceding examples and that yields new combination functions systematically.

2 The transformation framework

As explained in the Introduction, our viewpoint is that generation of combination functions is equivalent to the construction of suitable frame transformations mapping the range of the valuations into the extended real number line $[-\infty, +\infty]$. The intuition behind this viewpoint is that a combination operator is really just the standard arithmetic sum viewed through the warped glasses of the frame transformation. Mathematically, each admissible choice of a frame transformation induces a pullback to the valuation interval (which we will assume is $[-1, 1]$) of the standard vector space structure of the real numbers. Addition pulls back to a combination function and scalar multiplication pulls back to an operation which controls what we call the degree of skepticism of the members of the resulting family of combination functions. We develop the above concepts in the next few sections. We assume for simplicity that all valuations take values in the interval $[-1, 1]$. More general ranges of values can be dealt with by performing a straightforward preliminary symmetrization step as in [1].

2.1 Combination functions as nonlinear sums

We propose to consider as a combination function on the normalized measurement interval $[-1,+1]$ the binary operation $\oplus_\beta$ on $[-1,+1]$ that is conjugate to the standard addition operation $f(y_1,y_2) = y_2 + y_1$ on the interval $[-\infty,+\infty]$ via an appropriate *frame transformation* β from $[-1,+1]$ to $[-\infty,+\infty]$; we assume that β is an invertible and increasing map from $[-1,+1]$ onto $[-\infty,+\infty]$. In other words, we require that the diagram shown below be commutative, where $+$ denotes the usual arithmetic sum operator on $(-\infty,+\infty)$:

$$\begin{array}{ccc} (-1,+1)\times(-1,+1) & \xrightarrow{\beta} & (-\infty,+\infty)\times(-\infty,+\infty) \\ \Big\downarrow{\scriptstyle \oplus_\beta} & & \Big\downarrow{\scriptstyle +} \\ (-1,+1) & \xrightarrow{\beta} & (-\infty,+\infty) \end{array}$$

Equivalently, the combination function $\oplus_\beta$ on $[-1,+1]$ is defined by

$$a \oplus_\beta b = \beta^{-1}(\beta(a) + \beta(b)) \tag{11}$$

Visually, the frame transformation β deforms the standard valuation interval $[-1,+1]$ into the valuation range $[-\infty,+\infty]$. Each point x of $[-1,+1]$ is mapped to a corresponding point $\beta(x)$ of the interval $[-\infty,+\infty]$. Pairs of points are combined in $[-1,+1]$ in such a way that the result is the point that is mapped by the frame transformation β into the arithmetic sum of the images of these points. Different frame transformations define different deformations and thus lead to different combination functions, with the exception that frame transformations that are constant multiples of one another lead to the same combination function (c.f. the proof of Theorem 3.1).

2.2 Admissible frame transformations

We now consider the question of determining the properties that must be satisfied by a frame transformation $\beta : [-1,+1] \to [-\infty,+\infty]$ so that the combination function associated to β via conjugation as in Eq. 11 is admissible in the sense of Definition 1.1. Such a mapping β is called an *admissible frame transformation*.

Proposition 2.1. *A mapping $\beta : [-1,+1] \to [-\infty,+\infty]$ is admissible if and only if it is increasing and satisfies the boundary conditions $\beta(0) = 0$, $\beta(+1) = +\infty$.*

Proof. Recall the definition of $\oplus$ in terms of β from Eq. 11:

$$p \oplus q = \beta^{-1}(\beta(p) + \beta(q))$$

This definition assumes that β is an invertible mapping from $[-1,+1]$ to $[-\infty,+\infty]$. Thus, β must be either strictly increasing or strictly decreasing. We will now prove the necessary boundary conditions $\beta(0) = 0$, $\beta(1) = \infty$, which imply that β is increasing. Letting $q = 0$ above, we have:

$$\beta^{-1}(\beta(p)) = p = p \oplus 0 = \beta^{-1}(\beta(p) + \beta(0))$$

This equation holds for all values of p if and only if $\beta(0) = 0$. Next, let $q = 1$ above. Then we have:

$$\beta^{-1}(\beta(1)) = 1 = p \oplus 1 = \beta^{-1}(\beta(p) + \beta(1))$$

This equation is equivalent to $\beta(+1) = +\infty$. Thus, we have proved that admissibility is equivalent to the boundary conditions given in the statement of the Proposition.

It is straightforward to interpret the commutativity and belief / disbelief symmetry axioms for the corresponding combination function (as given following Definition 1.1) in terms of the frame transformation β, as we now show.

Proposition 2.2. *An admissible frame transformation $\beta : [-1, +1] \to [-\infty, +\infty]$ yields an associated combination function that satisfies the belief/disbelief symmetry property if and only if β has odd symmetry about* 0*, i.e.* $\beta(-x) = -\beta(x)$*.*

Proof. Let $q = -p$. Then assuming belief / disbelief symmetry and commutativity:

$$(p \oplus (-p)) = -((-p) \oplus p) = -(p \oplus (-p)),$$

so that

$$p \oplus (-p) = 0,$$

and therefore using the definition of $\oplus$:

$$\beta^{-1}(\beta(p) + \beta(-p)) = 0 \tag{12}$$

Applying β to both sides of this equation we see that

$$\beta(p) + \beta(-p) = \beta(0), \tag{13}$$

and letting $p = 0$ in particular it follows that

$$\beta(0) = 0$$

Eq. 13 now yields the desired conclusion that β has odd symmetry about 0. Conversely, if we know that β has odd symmetry about 0 then so does its inverse β^{-1}, and we see by Eq. 11 that the corresponding combination function $\oplus$ is commutative and exhibits belief / disbelief symmetry. This completes the proof of the Proposition.

2.3 The pulled-back metric

If one considers the subjective difference measure $\ominus$ as defined in Eq. 8, one may view the analog of Eq. 11 defining the combination function $\oplus$ via the frame transformation β as involving two distinct steps. In the first step, the pair (a, b) is mapped to the difference $\beta(a) - \beta(b)$, which is simply the *signed* version of the pullback to $[-1, 1]$ via (β, I_R) of the Euclidean metric on the real line R. In the second step, this signed distance function is pulled back to a metric on $[-1, +1]$ via β. Explicitly, the pulled-back metric referred to here is given by:

$$d(a, b) = |\beta(a) - \beta(b)| \tag{14}$$

Assuming that β is differentiable, the pulled-back metric is a Riemannian metric (see, e.g., [11]) on $[-1, +1]$ with length element ds given by:

$$ds = \beta'(x)dx \tag{15}$$

Observe that since $\beta(0) = 0$ by Proposition 2.1, the frame transformation β may be expressed in terms of the blown-up metric $ds = \beta'(x)dx$ quite simply:

$$\beta(x) = \int_0^x \beta'(u)du \tag{16}$$

Thus, the frame transformation β and the blown-up metric $\beta'(x)dx$ are completely equivalent: given either one of the two, the other can be constructed without difficulty. Since the frame transformation leads directly to the corresponding combination function, this implies that the combination function may also be constructed from the metric. In section 3 we will show that the metric may be constructed from the combination function (Theorem 3.1). Together with the above comments, this will show that the three basic objects of our theory, the combination function, the frame transformation, and the metric, are completely equivalent, so that if one of the three is specified then the other two may be constructed from it.

2.4 Nonlinear scaling and weighted combinations

Given a combination function $\oplus : [-1, +1] \to [-1, +1]$ obtained via a blow-up transformation $\beta : [-1, +1] \to [-\infty, +\infty]$, a new combination function is obtained by letting the group of scalings $x \mapsto tx$ for $t \in R^+$ act on $[-1, +1]$ via conjugation by the blow-up transformation β. Thus we have the commutative diagram shown below:

$$\begin{array}{ccc}
[-1, +1] & \xrightarrow{\beta} & [-\infty, +\infty] \\
\Big\downarrow{\scriptstyle \beta^{\leftarrow} t} & & \Big\downarrow{\scriptstyle x \mapsto tx} \\
[-1, +1] & \xrightarrow{\beta} & [-\infty, +\infty]
\end{array}$$

The collection of pullbacks $\beta^{\leftarrow} t$ forms a group of nonlinear scalings of the measurement interval $[-1, +1]$. The pulled-back scaling by t is given by:

$$(\beta^{\leftarrow} t)\, x = \beta^{-1}\left(t\beta(x)\right) \tag{17}$$

If we let the pulled-back scaling act on the combination function $\oplus_\beta$ on $[-1, +1]$ conjugate to the difference operator on $[-\infty, +\infty]$ via β, we obtain the following new combination function:

$$p \oplus_t q = \beta^{-1}\left(t\beta\left(\beta^{-1}(\beta(p) + \beta(q))\right)\right) = \beta^{-1}\left(t(\beta(p) + \beta(q))\right) \tag{18}$$

The corresponding blown-up metric on $[-1, +1]$ is:

$$ds = t\beta'(x)dx \tag{19}$$

We note that the scaled t-version of the combination function fails to be associative unless $t = 1$. Furthermore, it is not admissible in the sense of Definition 1.1, as it fails to satisfy the boundary conditions $p \oplus_t 0 = p, \quad p \oplus_t 1 = 1$. However, nonlinear scaling is quite useful in providing a scale of functions parametrized according to their *degree of skepticism*. By the latter term we are referring to the weight accorded to new information. This can be measured by comparing the a priori value p with the quantity

$$p \oplus_t 0 = \beta^{-1}\left(t\beta(p)\right) \tag{20}$$

which is the confidence assigned to the certainty level p by the combination function $\oplus_t$. The right-hand side of Eq. 20 is simply the result of scaling p by t as viewed in the frame defined by the transformation β (c.f. Eq. 17). We now define the degree of skepticism as the fraction of the confidence level that is rejected by the combination function $\oplus$. Notice that any combination function that satisfies the boundary condition $p \oplus 0 = p$ required for admissibility will automatically have a marginal skepticism of 0.

Definition 2.1. The *marginal skepticism* of a combination function $\oplus$ is the quantity

$$\sigma(\oplus) = \lim_{p \to 0} \left(1 - \frac{p \oplus 0}{p}\right) \tag{21}$$

The value of t determines the marginal skepticism of the scaled combination function $\oplus_t$ in the following very simple way.

Proposition 2.3.

$$\sigma(\oplus_t) = 1 - t \tag{22}$$

Proof. By definition of the marginal skepticism σ we have

$$\sigma(\oplus_t) = 1 - \frac{d}{dp}|_{p=0}\,(p \oplus_t 0)$$

It suffices to compute the derivative that appears on the right-hand side of this equation. Using the fact that $\beta(0) = 0$ (by Proposition 2.1) one finds:

$$\frac{d}{dp}|_{p=0}\left(\beta^{-1}(t\beta(p))\right) = \left(t\frac{\beta'(p)}{\beta'\left(\beta^{-1}(t\beta(p))\right)}\right)|_{p=0} = t$$

This concludes the proof of the Proposition.

Observe that the resulting skepticism in Proposition 2.3 is independent of the choice of frame transformation β. For $t = 1$ the skepticism is 0: confidence estimates are accepted at face value. Values of t greater than 1 yield negative skepticism, i.e. the combination function $\oplus_t$ amplifies confidence estimates. Values of t less than 1 yield skeptical combination functions that accept only a fractional portion of an incoming confidence estimate. We will say more below about the level of skepticism in connection with the rate at which consensus is attained in the presence of multiple sources of information.

Analogously, one may consider the nonlinear conjugated versions of linear combinations. In this way, one obtains operators such as the following:

$$p \oplus_{s,t} q = \beta^{-1}\left(s\beta(p) + t\beta(q)\right) \tag{23}$$

If the parameters s and t are chosen to satisfy $0 \le s \le 1$, $0 \le t \le 1$, $s + t = 1$, then Eq. 23 yields a nonlinear version of the convex combination operator $(x, y) \mapsto sx + ty$. Certain properties of convex combinations are shared by the nonlinear version. For example, one recovers one argument or the other as the parameter s approaches one of its limiting values:

$$p \oplus_{s,1-s} q \longrightarrow q \text{ as } s \to 0$$
$$p \oplus_{s,1-s} q \longrightarrow p \text{ as } s \to 1$$

Intermediate values of the weight parameter s yield other combinations of p and q; the closer s is to 0, the lower the weight accorded to p will be, while if s is close to 1 then p will be weighted more heavily than q in the combination. Notice that although this behavior is shared by the standard convex combination operators, the standard operators fail to satisfy the boundary conditions $0 \oplus q = q$ and $1 \oplus q = 1$. The new weighted nonlinear operators should be useful for purposes such as the combination of relevance ratings in information retrieval and the combination of preference ratings in recommendation systems (collaborative filtering). In such contexts the weights s and t may be used to give higher credence to certain information sources over others, based perhaps on prior experience.

2.5 Some admissible combination functions

Example 2.1 (Inverse hyperbolic tangent frame $\beta(x) = \tanh^{-1}(x)$). Using the fact that

$$\tanh^{-1}(x) = \frac{1}{2}\log\left(\frac{1+x}{1-x}\right) \tag{24}$$

we see that if we choose the frame transformation β to be the function $\tanh^{-1}$ in Eq. 11 then we obtain the following very simple expression for the associated combination function:

$$p \oplus q = \tanh\left(\tanh^{-1} p + \tanh^{-1} q\right) = \frac{p+q}{1+pq} \tag{25}$$

Nonlinear scaling by t as in Eq. 17 is given for the present choice of β by:

$$\left(\beta^{\leftarrow} t\right) x = \frac{(1+x)^t - (1-x)^t}{(1+x)^t + (1-x)^t} \tag{26}$$

and it follows from Eq. 24 and from the identity for the hyperbolic tangent of a sum contained in Eq. 25 that the combination function of Eq. 25 embeds as the case $t = 1$ of the family:

$$p \oplus_t q = \tanh\left(t\left(\tanh^{-1}(p) + \tanh^{-1}(q)\right)\right) = \frac{\left(\frac{1+p}{1-p}\right)^t - \left(\frac{1-q}{1+q}\right)^t}{\left(\frac{1+p}{1-p}\right)^t + \left(\frac{1-q}{1+q}\right)^t} \tag{27}$$

The pulled back Riemannian metric is given as in Eq. 15 by:

$$ds = t\left(\tanh^{-1}\right)'(x)dx = \frac{tdx}{1-x^2} \tag{28}$$

Nonlinear weighting leads to the operators

$$p\oplus_{s,t}q = \tanh\left(s\tanh^{-1}(p) + t\tanh^{-1}(q)\right) = \frac{\left(\frac{1+p}{1-p}\right)^s - \left(\frac{1-q}{1+q}\right)^t}{\left(\frac{1+p}{1-p}\right)^s + \left(\frac{1-q}{1+q}\right)^t} \tag{29}$$

As shown in [1], the inverse hyperbolic tangent frame transformation admits very interesting interpretations in terms of probability, Dempster-Shafer evidence theory, and the special theory of relativity.

Example 2.2 (Tangent frame $\beta(x) = \frac{2}{\pi}\tan(\frac{\pi}{2}x)$). This choice yields the following family of combination functions:

$$p\oplus_t q = \frac{2}{\pi}\tan^{-1}\left(\frac{t\sin(\frac{\pi}{2}(p+q))}{\cos(\frac{\pi}{2}p)\cos(\frac{\pi}{2}q)}\right) \tag{30}$$

The pulled back metric on the standard interval $[-1, +1]$ is:

$$ds = \frac{tdx}{\cos^2(\frac{\pi}{2}x)} \tag{31}$$

A significant difference between the tangent frame transformation considered here and the inverese hyperbolic tangent frame transformation of the preceding Example lies in their asymptotic behavior. The values of the hyperbolic tangent approach $+1$ exponentially fast as the argument approaches $+\infty$. On the other hand, the values of $2/\pi$ times the arctangent of y approach the limiting value $+1$ at the rate $1/y$ as $y \to +\infty$. We show below in Proposition 4.1 that this difference in asymptotic rates leads to a corresponding difference in the rates at which the combination functions based on these frame transformations aggregate values produced by a large number of source measures.

3 Recovering the frame transformation from the combination function

In this section we address the issue of determining whether a given combination function $\oplus$ that is admissible in the sense of Definition 1.1 is expressible via some frame transformation β as in Eq. 11. Our solution to this problem may be seen as a two-step process. We first show how to construct a special candidate frame transformation $\beta_\oplus$ directly from the original combination function $\oplus$. In order to determine whether $\oplus$ is transformation-based, one merely needs to check whether it is expressible in terms of this single special frame transformation $\beta_\oplus$. The second step of our process provides a method for checking whether $\oplus$ is expressible via $\beta_\oplus$.

Our frame transformation recovery process is useful from a practical point of view since it provides an explicit method for constructing a frame transformation that yields a given combination function. With such a frame transformation in hand, one may proceed to generalize the original combination function by using nonlinear scaling operations as described in the preceding sections. Furthermore, our results are interesting from a theoretical viewpoint, as they show the equivalence of three basic objects of our theory: the combination function $\oplus$, the frame transformation β, and the blown-up metric $\beta'(x)dx$. A frame transformation β may easily be expressed in terms of the corresponding blown-up metric $ds = \beta'(x)dx$ as in Eq. 16. The following result shows that the metric $\beta'(x)dx$ and the frame transformation β may be recovered (modulo a scale factor) from the combination function $\oplus$.

Theorem 3.1. *Let $(p, q) \mapsto p \oplus q$ be a continuously differentiable combination operator such that $p \oplus z(p) = 0$ for some function $p \mapsto z(p)$. Let β denote an arbitrary frame transformation. Then the following statements are equivalent:*

1. *$\oplus$ is conjugate to the arithmetic sum operator via the frame transformation β*

2. *β is of the form $C\beta_\oplus$, where C is a nonzero constant and $\beta_\oplus$ is the special frame transformation defined by:*

$$\beta_\oplus(p) = \int_0^p (\partial_1 \oplus)(x, z(x))\ dx \tag{32}$$

3. *$\oplus$ is conjugate to the arithmetic sum operator via the frame transformation $\beta_\oplus$ as given in Eq. 32*

4. *The composite function $\phi := \beta_\oplus \circ \oplus$ satisfies the partial differential equation*

$$\partial_1 \partial_2 \phi = 0 \tag{33}$$

Proof.

- ((1) implies (2)): If $\oplus$ is conjugate to $+$ via β, then we must have:

$$\beta(p \oplus q) = \beta(p) + \beta(q) \tag{34}$$

  Taking partial derivatives with respect to p we obtain:

$$\beta'(p \oplus q)(\partial_1 \oplus)(p, q) = \beta'(p) \tag{35}$$

  Letting $q = z(p)$, and observing that $p \oplus z(p) = 0$ we have:

$$\beta'(0)(\partial_1 \oplus)(p, z(p)) = \beta'(p) \tag{36}$$

Therefore:

$$\beta'(p) = \beta'(0)(\partial_1 \oplus)(p, z(p)) \tag{37}$$

Integration w.r.t. p now yields $\beta = C\beta_0$, with $C = \beta'(0)$:

$$\beta(p) = \beta'(0) \int_0^p (\partial_1 \oplus)(x, z(x))\ dx \tag{38}$$

This proves that (2) holds.

- ((2) implies (3)): Just observe that the conjugacy condition is invariant under scalings. That is, if one assumes that $\oplus$ is conjugate to $+$ via β:

  $$p \oplus q = \beta^{-1}\left(\beta(p) + \beta(q)\right),$$

  and if K is any nonzero constant, then since the inverse of the scaled transformation $K\beta$ is given by

  $$(K\beta)^{-1}(y) = \beta^{-1}\left(\frac{y}{K}\right),$$

  the fact that multiplication by K distributes over addition yields

  $$p \oplus q = (K\beta)^{-1}\left((K\beta)(p) + (K\beta)(q)\right),$$

  so that $\oplus$ is also conjugate to $+$ via the scaled transformation $K\beta$. Choosing $K = 1/C$, one now obtains (3) from (2).

- ((3) implies (4)): If (3) holds, then $\phi(p, q) := \beta_0(p \oplus q) = \beta_0(p) + \beta_0(q)$ clearly satisfies $\partial_1\partial_2\phi = 0$.

- ((4) implies (1)): Suppose that (4) holds. Then iterated partial integration shows that there exist functions α_1 and α_2 such that

  $$\beta_\oplus(p \oplus q) = \phi(p, q) = \alpha_1(p) + \alpha_2(q)$$

  Notice that the functions α_i are defined only up to an additive constant. Assume without loss of generality that $\alpha_2(0) = 0$. Since $p \oplus 0 = p$, we obtain

  $$\beta_\oplus(p) = \phi(p, 0) = \alpha_1(p) + \alpha_2(0) = \alpha_1(p)$$

  Now use the fact that $0 \oplus q = q$. We must have

  $$\begin{aligned}\beta_\oplus(q) &= \phi(0, q) = \alpha_1(0) + \alpha_2(q)\\ &= \beta_\oplus(0) + \alpha_2(q) = \alpha_2(q)\end{aligned}$$

  We have now shown that both α_1 and α_2 equal $\beta_\oplus$, so that

  $$\beta(p \oplus q) = \phi(p, q) = \beta_\oplus(p) + \beta_\oplus(q)$$

  We know that $\beta_\oplus$ is strictly increasing and therefore invertible on its image, so we conclude that $\oplus$ is conjugate to the standard arithmetic sum operator $+$ via the frame transformation $\beta_\oplus$. This proves (2) and (1), and concludes the proof of the Theorem.

Example 4. Let us now re-examine the probabilistic combination operator considered in the Example of the Introduction in the light of the above Theorem. Recall the form of the combination operator:

$$p \oplus q = p + q - pq$$

Notice that $p \oplus \frac{p}{p-1} = 0$, i.e. we have $z(p) = \frac{p}{p-1}$. The frame transformation β must satisfy:

$$\beta(p + q - pq) = \beta(p) + \beta(q)$$

Differentiating with respect to p and letting $q = \frac{p}{p-1}$ we have:

$$\beta'(0)\frac{1}{1-p} = \beta'(p)$$

As in the proof of the Theorem we may assume that $\beta'(0) = 1$, so that the metric is given by

$$\beta'(p) = \frac{1}{1-p}$$

and we obtain the frame transformation

$$\beta(p) = \log\left(\frac{1}{1-p}\right)$$

as given in the Example of the Introduction. A straightforward calculation shows that the composite operator $\phi := \beta \circ \oplus$ satisfies the partial differential equation given in the Theorem. Indeed, we have

$$\phi(p, q) = -\log(1 - p - q + pq)$$

so that

$$\partial_p \phi(p, q) = \frac{1-q}{1 - p - q + pq}$$

and therefore

$$\partial_q \partial_p \phi(p, q) = \frac{1 - p - q + pq - (1-q)(1-p)}{(1 - p - q + pq)^2} = 0$$

as claimed. The implication (3) $\Rightarrow$ (2) of the Theorem thus implies that we have found a correct conjugating frame transformation β. Of course, even before checking that ϕ satisfies the partial differential equation one may have noticed that $\phi(p, q)$ may be decomposed as follows:

$$\phi(p, q) = -\log((1-p)(1-q)) = \log\left(\frac{1}{1-p}\right) + \log\left(\frac{1}{1-q}\right)$$

which is merely a restatement of the conjugacy condition itself. Nonetheless, in more involved examples it may be difficult to see that the analogous decomposition holds for ϕ in such a direct fashion; in such cases it is advantageous to apply the partial differential equation criterion as was done above.

Physical Interpretation of Theorem 3.1

The characterization in Theorem 3.1 of the composition $\phi = \beta \circ \oplus$ as the solution of a partial differential equation may be interpreted in physical terms related to wave propagation. Defining new variables (x, y) from the variables (p, q) by:

$$x = p + q$$
$$y = p - q$$

the partial differential condition on $\widetilde{\phi}(x, y) := \phi(p, q)$ becomes

$$\frac{\partial}{\partial x^2}\widetilde{\phi} = \frac{\partial}{\partial y^2}\widetilde{\phi} \tag{39}$$

which is the classical equation describing linear wave propagation [17]. The new variables x and y may be interpreted as space and time. The old variables p and q represent position as viewed in frames moving at the wave velocity in opposite directions. Taking into account the restriction that in the old variables

$$\phi(p, z(p)) = 0,$$

we have in the new variables

$$\widetilde{\phi}(x, y)|_{x-y=z(x+y)} = 0 \tag{40}$$

Equations 39 and 40 together constitute a so-called boundary-value problem. As is known from the theory of partial differential equations, an additional boundary condition should be specified in order for the boundary-value problem to be uniquely solvable. For example, information concerning the rate of change of the function ϕ in a direction transverse to the "zero curve" $q = z(p)$ (or $x - y = z(x + y)$) would be sufficient. In any case, we see that the function ϕ may be constructed by specifying "initial data" on the curve $q = z(p)$ and allowing this information to "propagate" via the wave equation (Eq. 39).

4 Asymptotic consensus growth

In this section we address the growth of the degree of consensus in the presence of multiple sources of information. We assume that an infinite sequence of observations, each having certainty value p, is provided to a system that uses a combination function $\oplus$ to aggregate certainty values. The main issue is to quantitatively describe the aggregation of certainty as the number of observations increases without bound. The issue of the asymptotic consensus growth *rate* is an important one. For example, the creators of MYCIN encountered difficulties associated with the fact that their combination function leads to very rapid growth of consensus [3]. We will show that our framework allows the growth rate to be controlled by choosing appropriate frame transformations. We will also show that the degree of skepticism of nonlinearly scaled transformation-based combination functions is reflected in the asymptotic consensus *value* as the number of sources increases.

Concretely, the situation at hand is as follows. Given a combination function $\oplus$ and given a number p between 0 and 1, consider the sequence $(p_n)_{n\in\mathrm{N}}$ defined by:

$$\begin{aligned} p_0 &= 0 \\ p_{n+1} &= p_n \oplus p \end{aligned} \tag{41}$$

In words, p_n is the combined degree of certainty associated with n certainty judgements of value p, according to the combination function $\oplus$. We are interested in determining the behavior of p_n for large values of n.

4.1 Admissible transformation-based combination functions

Let us begin by illustrating the sort of analysis that we are interested in, for the special case of the probabilistic combination function given in the Example of the Introduction. In this case one obtains the following sequence of combined certainty estimates as in Eq. 41:

$$\begin{aligned} p_0 &= 0 \\ p_{n+1} &= p_n + p - p_n p = p_n(1-p) + p \end{aligned} \tag{42}$$

The p_n are therefore the partial sums of a geometric sequence:

$$p_n = p \sum_{j=0}^{n-1} (1-p)^j = 1 - (1-p)^n \tag{43}$$

and approach the limiting value 1 exponentially fast as $n \to 0$. Our analysis in terms of frame transformations below will show that this rate of convergence follows from the asymptotic behavior of the inverse frame transformation in this case.

Proposition 4.1. *Let $\oplus$ be an admissible transformation-based combination function with associated frame transformation β. Define the sequence (p_n) of combined values as in Eq. 41. Then*

$$p_n \to 1 \quad \textit{as} \quad n \to \infty \tag{44}$$

Furthermore, convergence occurs at the rate

$$p_n = \beta^{-1}(Cn) \tag{45}$$

with $C = \beta(p)$, where p is the confidence value that generates the sequence (p_n).

Proof. Start with a combination function based on a frame transformation β as in Eq. 11:

$$a \oplus_\beta b = \beta^{-1}(\beta(a) + \beta(b)) \tag{46}$$

The sequence of combined certainty estimates defined in Eq. 41 becomes:

$$\begin{aligned} p_0 &= 0 \\ p_{n+1} &= \beta^{-1}(p_n + p) \end{aligned} \tag{47}$$

Define:

$$\pi_n = \beta(p_n), \qquad \pi = \beta(p) \tag{48}$$

Then one has

$$\begin{aligned} \pi_0 &= 0 \\ \pi_{n+1} &= \pi_n + \pi \end{aligned} \tag{49}$$

Therefore:

$$\pi_n = n\pi, \tag{50}$$

so that the p_n approach $\beta^{-1}(\infty) = 1$ as $n \to \infty$. The rate of consensus growth is determined by the asymptotic behavior of the frame transformation β. Indeed, by Eq. 50 one obtains

$$p_n = \beta^{-1}\left(n\beta(p)\right)$$

This completes the proof of the Proposition.

The preceding Proposition shows that if $\beta^{-1}(x)$ approaches 1 exponentially fast as $x \to \infty$, as is the case for the probabilistic combination function considered above, for which $\beta^{-1}(y) = 1 - e^{-y}$, then p_n also approaches 1 exponentially fast as $n \to \infty$. Other growth rates translate from β^{-1} to the sequence p_n analogously. For example, the tangent frame transformation yields a sequence p_n that approaches 1 like $1/n$. This provides the ability to control the asymptotic consensus growth rate, thus offering a way to avoid the problems encountered with the MYCIN combination function.

4.2 Skeptical combination functions

Next we are interested in studying the nature of consensus growth for skeptical transformation-based combination functions. Specifically, consider the combination function $\oplus_t$ corresponding to the frame transformation β including nonlinear scaling by t as in Eq. 18:

$$p \oplus_t q = \beta^{-1}\left(t\beta\left(\beta^{-1}(\beta(p) + \beta(q))\right)\right) = \beta^{-1}\left(t(\beta(p) + \beta(q))\right) \tag{51}$$

In particular, if $\beta = \tanh^{-1}$ then one has the combination function

$$p \oplus_t q = \frac{\left(\frac{1+p}{1-p}\right)^t - \left(\frac{1-q}{1+q}\right)^t}{\left(\frac{1+p}{1-p}\right)^t + \left(\frac{1-q}{1+q}\right)^t} \tag{52}$$

The parameter t is a positive number but is otherwise free. If $t = 1$, this combination function is rather similar to the MYCIN combination function of Eq. 2:

$$p \oplus_1 q = \frac{\left(\frac{1+p}{1-p}\right) - \left(\frac{1-q}{1+q}\right)}{\left(\frac{1+p}{1-p}\right) + \left(\frac{1-q}{1+q}\right)} = \frac{p+q}{1+pq} \tag{53}$$

It was shown above that, regardless of the choice of frame transformation β, the nonlinearly scaled operator $\oplus_t$ exhibits skeptical behavior when $t < 1$. We measure the degree of skepticism using the notion of marginal skepticism; we showed that the marginal skepticism of $\oplus_t$ is $1 - t$.

We study the convergence of the sequence p_n associated by the combination function $\oplus_t$ to a collection of n judgements of certainty p. With notation as above we have:

$$\begin{aligned} p_0 &= 0 \\ p_{n+1} &= \beta^{-1}\left(t\left(\beta(p_n) + \beta(p)\right)\right) \end{aligned} \tag{54}$$

In contrast to the case of admissible combination functions discussed above, for the nonlinearly scaled combination functions $\oplus_t$ with $t < 1$, the rate of convergence of the p_n toward their limiting value is always exponential. However, the limiting *value* p_∞ depends on the scaling parameter t and may therefore be controlled.

Proposition 4.2. *Let $\oplus$ be an admissible transformation-based combination function with associated frame transformation β. Consider the sequence (p_n) defined in terms of the skeptical t-version $\oplus_t$ of $\oplus$ as in Eq. 54. Then*

$$p_n \to \beta^{-1}(C\beta(p)) \quad as \quad n \to \infty,$$

where $C = t/(1-t)$. In particular, if $t < 1$ then the limiting value is strictly less than 1. The rate of convergence is exponential whenever $t < 1$.

Proof. Define

$$\pi_n = \beta(p_n), \qquad \pi = \beta(p) \tag{55}$$

Then one has

$$\begin{aligned} \pi_0 &= 0 \\ \pi_{n+1} &= t\left(\pi_n + \pi\right) \end{aligned} \tag{56}$$

If $t = 1$, one then sees that $\pi_n = n\pi$, so that the p_n approach $\beta^{-1}(\infty) = 1$ as $n \to \infty$ as described above in our analysis for admissible combination functions. In the case $t < 1$, the linear recurrence in Eq. 56 may be solved by using the method of variation of constants, yielding:

$$\pi_n = \sum_{j=0}^{n-1} t^{n-j} t\pi = \frac{t\pi}{1-t}\left(1 - t^n\right), \tag{57}$$

Eq. 57 shows that the rate of convergence toward the limiting value is always exponential in the case $t < 1$. It also follows in the case $t < 1$ that the limiting value as $n \to \infty$ is:

$$\pi_\infty = \frac{t\pi}{1-t} \tag{58}$$

The limiting value of the p_n is now obtained from Eq. 58 by using Eq. 55:

$$p_\infty = \beta^{-1}(\pi_\infty) = \beta^{-1}\left(\frac{t\pi}{1-t}\right)$$

This completes the proof.

Proposition 4.2 shows that the asymptotic limit p_∞ of the p_n is obtained from the "seed" value p by a nonlinear scaling transformation with steepness parameter $t/(1-t)$. The limit p_∞ of the p_n is the inverse image via β of the finite number on the right-hand side of Eq. 58 and is thus strictly less than 1. For example, if $t = 1/2$ one has $p_\infty = p$. Values of t greater than 1/2 yield values of p_∞ between p and 1, while values of t smaller than 1/2 yield values of p_∞ less than p, which is "skeptical" behavior. An asymptotic version of the degree of marginal skepticism of Definition 2.1 may be defined here in a natural way:

$$\sigma_\infty = 1 - \lim_{p \to 0} \frac{p_\infty}{p}$$

It is easy to see that the asymptotic marginal skepticism σ_∞ is given here by:

$$\sigma_\infty = 1 - \frac{t}{1-t}$$

In terms of the marginal skepticism $\sigma = 1 - t$ of the combination function $\oplus_t$, one has:

$$\sigma_\infty = 2 - \frac{1}{\sigma}$$

Thus, the asymptotic marginal skepticism is an increasing function of the marginal skepticism of the underlying combination function $\oplus_t$.

Conclusions

We have presented a new framework which provides a unified foundation for the construction of combination operators for use in such areas as confidence aggregation in knowledge-based systems, relevance rating combination in information retrieval, and lateralization assessment in neurobiology. Our framework is based on the postulate that different combination operators are warped versions of the standard arithmetic sum operator as viewed in appropriate frames of reference. We have given examples showing that certain probabilistic combination operators and MYCIN-like combination operators arise in this way. In addition to unifying such previously considered operators, our framework provides a nonlinear scaling mechanism that allows one to modify a given combination operator by providing parametrized families of operators that extend the original operator. We have shown that this feature allows control over the degree of skepticism of the operators, i.e. their sensitivity to new information. We provide an algorithmic method to check whether a given combination operator fits into our framework or not, and that constructs an appropriate reference frame relating the operator to the arithmetic sum operator whenever such a frame exists. Furthermore, we have shown that our framework makes it easy to construct new combination operators, merely by selecting among the infinitely many admissible frame transformations available. Finally, we have shown that our framework provides control over the rate at which the combined measure increases when combining a large number of source measures. This should allow one to address the difficulties associated with excessively high convergence rates such as those produced by the ad-hoc combination operator used in the classical knowledge-based system MYCIN.

References

[1] S.A. Alvarez. "Rational comparison of probabilities via a blow-up conjugacy", Technical Report No. 97-NA-010, Center for Nonlinear Analysis, Carnegie Mellon University, Aug. 1997

[2] S.A. Alvarez, S.L. Levitan, J.A. Reggia. "Metrics for Cortical Map Organization and Lateralization", *Bulletin of Mathematical Biology*, vol. 60 (1998), 27-47

[3] B.G. Buchanan. Personal communication

[4] B.G. Buchanan, E.H. Shortliffe (eds.). *Rule-Based Expert Systems: the MYCIN experiments of the Stanford heuristic programming project*, Addison-Wesley, 1984

[5] R.J. Davidson, K. Hugdahl. *Brain Asymmetry*, MIT Press, 1995

[6] L.B. Day, P.F. MacNeilage. "Postural Asymmetries and Language Lateralization in Humans (*Homo sapiens*)", *Journal of Comparative Psychology*, vol. 110 (1996), no. 1, 88-96

[7] G. Giacomelli, R. Giacomelli. "Results from Accelerator Experiments: The Left-Right Asymmetry at SLAC", http://axoph1.cern.ch/papers/results_from_acc_exp/results_from_acc_exp.html

[8] W.D. Hopkins. "Hand Preferences for a Coordinated Bimanual Task in 110 Chimpanzees (*Pan troglodytes*): Cross-Sectional Analysis", *Journal of Comparative Psychology*, vol. 109 (1995), no. 3, 291-297

[9] S.L. Lauritzen and P.P. Shenoy. "Computing marginals using local computation", preprint

[10] P. Maruff, D. Hay, V. Malone, J. Currie. "Asymmetries in the Covert Orienting of Visual Spatial Attention in Schizophrenia", *Neuropsychologia*, vol. 13 (1995), no. 10, 1205-1223

[11] B. O'Neill. *Semi-Riemannian geometry with applications to relativity*, Pure and Applied Mathematics, vol. 103, Academic Press, 1983

[12] A.R. Palmer, C. Strobeck. "Fluctuating asymmetry as a measure of developmental stability: Implications of non-normal distributions and power of statistical tests", *Acta Zoologica Fennica*, vol. 191 (1992), 55-70
Also see http://gause.biology.ualberta.ca/palmer.hp/pubs/92P+S/92P+S.htm

[13] Heckerman. "Probabilistic interpretation for MYCIN's certainty factors", in *Uncertainty in Artificial Intelligence*, L.N. Kanal and J.F. Lemmer, eds., Elsevier / North-Holland, 1986

[14] S. Russell, P. Norvig. *Artificial Intelligence: a modern approach*, Prentice-Hall, 1995

[15] G. Shafer. *A Mathematical Theory of Evidence*, Princeton University Press, 1976

[16] P.P. Shenoy and G. Shafer. "Axioms for probability and belief-function propagation", in *Uncertainty in Artificial Intelligence*, 4 (1990), 169-198.

[17] W. Strauss. *Partial Differential Equations: An introduction*, Wiley, 1992